N71-22478

NASA CR-117842

EARTH-PHYSICS DATA-MANAGEMENT STUDY

Final Report

Grant NGR 09-015-107

January 1971

Prepared for

National Aeronautics and Space Administration
Washington, D.C. 20546

Smithsonian Institution
Astrophysical Observatory
Cambridge, Massachusetts 02138

EARTH-PHYSICS DATA-MANAGEMENT STUDY

Final Report

Grant NGR 09-015-107

January 1971

Prepared for

National Aeronautics and Space Administration
Washington, D.C. 20546

Smithsonian Institution
Astrophysical Observatory
Cambridge, Massachusetts 02138

101-15

TABLE OF CONTENTS

EARTH-PHYSICS DATA-MANAGEMENT STUDY

Final Report

1. INTRODUCTION, THE CONCEPT OF AN EARTH-PHYSICS INFORMATION-MANAGEMENT SERVICE

Space techniques provide powerful methods to measure physical quantities involved in the phenomena of the earth as a mechanical system (Kaula, 1970). Results to date and the promise of future advances justify a vigorous earth-physics program within the national space effort. Such a program recognizes that mechanical properties and processes of the earth are a significant element in a comprehensive description and understanding of the environment in which man must live.

Studies such as that at Williams College during the summer of 1969 have discussed the instrumentation and procedures that will generate a broad spectrum of new information about the earth (Kaula, 1970). Less attention has been directed to the problems associated with organization, accessibility, and application of the information generated. This study addresses some aspects of data-information management for an earth-physics program.

The information-management concept envisioned here would usually not be concerned with primary observational data. These data would presumably be handled by appropriate data banks, following current practice. The World Data Centers that have evolved from IGY and IQSY activities are examples of such existing capabilities. For the United States, the National Space Science Data Center (NSSDC) at Goddard Space Flight Center (GSFC) is another example.

Similarly, the many specialized scientific groups that employ the primary observations to derive fundamental earth-physics results would be expected to continue their necessary roles. They would publish or otherwise make available the resultant information in their preferred medium. The International Polar Motion Service (IPMS), the Bureau Internationale de l'Heure (BIH), and the United States Naval Observatory (USNO) are examples of organizations of this type.

From such primary sources, the information-management activity considered here would assemble and transcribe in computer-accessible form an appropriate selection of information to support an earth-physics applications program. For convenience in this report, the name Earth-Physics Information-Management Service (EPIMS) will denote this activity.

The available information about the earth, particularly that generated by space techniques, has by 1970 become very detailed. The involved formalism required to represent such detail dictates that many applications of this information must utilize a sizable computer. This is to be expected because the fundamental information in almost all instances has been achieved by sophisticated computer operations. Thus, it is natural to inquire whether there is some efficient technique, still utilizing computer methods, that can be used to manage the accumulated information and make it conveniently available to all potential users. In this age, the printed page would not seem to be an efficient or optimum intermediary between computer-generated information about the earth and computer-formulated applications of this information.

Many computer-oriented information-management approaches could be conceived, of course. One intriguing possibility would offer interactive, remote-terminal access to a major computer in which the information would be retained in a convenient manner. If the peripheral equipment at the computer were sufficiently versatile, any organization with a terminal for computer access via telephone circuits could also call upon the earth-physics information service. Many, and probably most, potential user organizations already have such terminals. Billing for computer time could be handled in the way that remote-terminal computing is currently being billed.

The interactive feature of the system would allow an unfamiliar user to inquire as to what information is available and in what format, and how to reach it. Using his remote terminal, the user could write his own program to transform the information into the form he requires, and perhaps to perform desired calculations based on the data.

The specifications that must be met by such an information-management system follow first from the nature, format, and scope of the available and anticipated information. Second, the specifications must reflect an assessment or judgment as to which data will be in sufficient demand to justify treatment by the envisioned techniques. Topic by topic, Section 2 of this report addresses these aspects of the information-management problem.

In organizing the conclusions from individual considerations of these discipline topics, a distinction is made between a data base of fundamental information — usually a table of numbers — and a collection of useful standard processes to be performed using the data base — in the form of a collection of computer programs. One or more sets of spherical-harmonic coefficients for the gravitational potential of the earth would be a typical segment of the data base. Finding the potential at a required point on the earth is an example of a typical process employing the data base.

Subsequent tables collect a number of candidate data sets and candidate processes for inclusion in EPIMS. The principal objective in so presenting these sets and processes is to assemble the necessary statistics on computer requirements, so that the size of the data-management problem can be expressed quantitatively.

Various options for the scope of a data-management activity can be assembled from subsets of the candidate data sets and processes. These options, with their different computer requirements, will influence the magnitude and organization of the computer hardware and software required to implement a data-management system of the sort suggested.

Section 3 of this report considers the general problems involved in organizing a computer system for managing the data from an earth-physics program. A number of significant questions surround such an operation.

Finally, Section 4 presents an initial conclusion about the practicality and advisability of a computer-oriented information-management service for earth physics. Recommendations are also given for appropriate next steps if an EPIMS is to be implemented.

2. DATA BASES AND PROCESSES FOR EARTH-PHYSICS DISCIPLINES

2.1 Definitions

To begin a review of what items should be included in an information-management activity for an earth-physics program, a definition delineating the content of the program must first be adopted. The present study adopts the definition set forth in the Williamstown Report (Kaula, 1970).

The following sections contain discussions of scientific disciplines that enter an earth-physics program as envisioned in the Williamstown study. Each section briefly summarizes the status of computer-compatible models that might be included in a data-management plan.

To facilitate understanding the tables that accompany the sections on the various scientific disciplines, two definitions may be helpful.

The first deals with the differentiation between the data or information base and the concept of a process. The terms "data" and "information" will sometimes be used interchangeably here, although in general, "data" refers to the raw material from which information is derived. An information base is loosely defined as a collection of nonexecutable storage bits that can be accessed by a process in order that a higher level of knowledge can be achieved. A process is a computer program (or collection of programs) that performs a prescribed transformation on a specified information base. A process may or may not carry an information base with it.

The second definition deals with the notion of a character of storage. For purposes herein, a character of storage on some storage device is represented by eight consecutive binary digits capable of being coded into a representation of the usual set of alphanumeric characters such as the decimal digits, letters of the alphabet, and other special symbols.

2.2 Time and Frequency Standards

An accurate knowledge of time and frequency is fundamental to many of the earth-physics areas, particularly to space activities that provide data for earth physics. Several agencies around the world, especially national time services, maintain atomic oscillators and clocks to provide fundamental time references. Various earth-physics measurements, therefore, may relate epoch and frequency to one or another of these time services. For comprehensive global analyses, these measurements must be brought to a common time system. Thus, it will be essential that the EPIMS have available the relations between the various atomic times.

The several time services transmit radio signals that carry time and frequency information. These signals are received and used in various ways by observation stations taking data for earth-physics objectives. Knowledge of the relations between the various transmitted signals is necessary in order to correct the observational data to a common reference.

Until the development of atomic clocks, fundamental time was determined from astronomical measurements. Time derived from orbital motions of bodies in the solar system is tabulated as "ephemeris time." Since the motions of bodies in the solar system — for example, the moon — are involved in earth-physics measurements, the information-management service must provide the relations between atomic time and the independent variable in pertinent ephemerides.

Table 1 lists the information on time and frequency standards that should be considered for incorporation in the EPIMS.

2.3 Earth's Rotational Motion

Historically, measurements of the rotational motion of the earth have been closely associated with time determinations. In fact, tabulations of the rotational position of the earth about its axis are still referred to as "universal time" of one sort or another, depending on what corrections have been applied.

Table 1. Information for time and frequency standards.

Service	Description	Size (characters)	Classification
Relation between time services	Provides the relations between the atomic times, e.g., BIH, USNO, NBS.	10^5	Base
Relation between frequency standards	Provides the rate relationship between the various frequency standards used by the time services.	10^5	Base
Relation between transmitted signals	Provides the measured difference between approximately 12 time and frequency transmitting stations and their governing standards.	10^5	Base
Transmission-time correction	Provides values from model of radio-propagation interval between time-signal transmitter and receiver at arbitrary coordinates.	10^5	Process
Ephemeris time	Provides the relations between atomic time and the ephemeris time associated with various ephemerides for the solar system.	10^5	Base

The relations between the various universal times tabulated by several agencies and atomic time must be available (see Table 2).

Several agencies currently determine the motions of the earth's axis of rotation. These individual determinations should be made available, as well as some best fitting mean set (see Table 2). Preliminary requests from potential users of such information are likely to be concerned with what type of data are available for a specified time, what is the interval between successive determinations, and what is the estimated accuracy. The request for the actual data may be for discrete values determined by a specific agency or agencies or for interpolated instants.

2.4 Positional Astronomy

In the previous section, the existence of two reference coordinate systems was implicitly assumed. One is a fundamental set of three orthogonal directions defined in terms of the directions to distant astronomical sources. The second is a set of orthogonal axes in the earth with its origin at the center of mass of the earth. The former is essentially an inertial system and is a product of positional astronomy. The latter comes from geometrical geodesy (as discussed in the next section). The rotational motions of the earth are precisely specified in terms of the angles relating the directions of these two fundamental systems.

Positional astronomy has several data collections that are appropriate for inclusion in the information-management system (see Table 3). The most obvious and perhaps the most easily initiated data bases are the star catalogs. Primarily, the FK4 catalog (Fricke and Kopff, 1963) and the SAO Star Catalog (1966) (in the FK4 system) should be made accessible, although it is possible that special interest would exist for other types of catalogs. A radio star catalog should be maintained and updated as new observational data accumulate. The radio star catalog is especially important to very long-baseline interferometry measurements.

Ephemerides of the planets, the moon, and perhaps other members of the solar system would be useful to some aspects of an earth-physics program.

Table 2. Information for earth's rotational motions.

Service	Description	Size (characters)	Classification
Rotation of the earth	Provides relations between the various determinations of UT1, related universal times, and atomic time.	10^4	Base
Polar motion	Provides relations between the various determinations of pole position.	10^5	Base
Models of rotation of the earth	Provides an analytical model for features of the rotation of the earth, such as expected annual variations.	10^5	Process
Observatory list for the earth's rotation	Lists coordinates and other characteristics of the observatories participating in measurements of the earth's rotational motions.	10^4	Base

Table 3. Information for positional astronomy.

Service	Description	Size (characters)	Classification
SAO Star Catalog	Provides identification, right ascension, and declination position coordinates and proper motion for over 259,000 stellar objects covering the whole sky. Additional information concerning spectral and visual classifications is also provided.	10^7	Base
Bonner Durchmusterung (BD) and Cordoba Durchmusterung (CD) Catalogues	Provide stellar-position information on 473,000 and 706,000 objects, respectively, covering virtually the whole sky.	10^8	Base
Discrete Radio Source Catalog	Provides information on approximately 26,000 radio sources, including position, flux, spectral slope, polarization, and optical corollation. Such information would be useful to very long-baseline projects. A subset of this catalog would be those radio sources of small angular size that can be used as a fundamental reference system.	10^6	Base
General tables of the American Ephemeris and Nautical Almanac	Provides a number of collective tables – e.g., calendar of phenomena, universal and sidereal times, ephemeris of the moon, and pole-star movements.	10^5	Base
Differential precession and nutation	Provides the necessary values of the correction for differential precession and nutation to be added to the observed difference in right ascension and declination of an object relative to a comparison star.	10^4	Process
Differential aberration	Provides corrections for differential stellar aberration similar to those used for precession and nutation.	10^4	Process
Lunar ephemeris	Represents lunar ephemeris either from general perturbation treatment or as numerical integration with parameters adjusted to best observational data.	10^5	Process
Lunar libration	Represents lunar librations with parameters adjusted to best observational data.	10^5	Process

Planetary ephemerides could be provided as in the American Ephemeris and Nautical Almanac or as the product of an integration of planetary motions produced in connection with work done for the space program. Similarly, the lunar orbit and librations should be best obtained from analyses of data generated by techniques such as laser ranging.

2.5 Geometrical Geodesy

In current practice, axis directions for a coordinate system for the solid earth are defined by adopted positions for the observatories measuring the rotational motions of the earth. The origin for this system is the center of mass of the earth, which can be determined from satellite orbits. The distance scale is specified operationally by the adopted value of the speed of light.

The information-management system should store all the pertinent information that enters into the detailed specification of a terrestrial coordinate system (Table 4). This will then be the system in which coordinates are tabulated for all the observing instruments involved in the earth-physics program. It will also be useful to have available the transformations that relate the classical survey datums to the fundamental terrestrial system. Many of these data are already best determined from analyses of satellite observations, and further improvements in them will be a major product of the earth-physics program.

Eventually, the whole concept of a fixed terrestrial coordinate system must be reexamined, particularly when the accuracy of the coordinate specification approaches the scale of annual crustal-plate motions. This need not be included in initial versions of the information-management system.

2.6 Earth's Gravitational Potential

The geopotential is most usefully represented currently by means of spherical-harmonic expansions. The definition of the expansion used should

Table 4. Information for geometrical geodesy.

Service	Description	Size (characters)	Classification
Fundamental geocentric coordinate system	Gives instrument coordinates that provide an implicit defintion of the fundamental geocentric coordinate system.	10^5	Base
Directory of geodetic control points	Gives fundamental coordinates of principal points in the several geodetic datums and information on how they were surveyed or determined. Includes approximately 5000 entries.	10^5	Base
Datum transformations	Provides transformations between the several geodetic survey datums and the fundamental geocentric coordinate system defined with the aid of satellite geodesy.	10^3	Process
Physiography	Describes physical characteristics of the surface of the earth on a $1° \times 1°$ element, giving such information as the average, highest, and lowest elevations, rock type, and temperature ranges.	10^6	Base
Satellites for simultaneous observations	Lists the characteristics of satellites appropriate for simultaneous observations for geometrical satellite geodesy.	10^5	Base
Earth tides	Provides a model of physical motion of points on the earth's surface caused by earth tides.	10^6	Process

be clearly stated, and facilities should exist for converting the coefficients to one or another of the principal methods of normalization. Degrees of the harmonics up to about 25, or about 650 coefficients, can be expected in the near future. The geopotential or functions thereof may be requested for specific points, and programs should be made available to compute the following (see Table 5):

A. Free-air gravity anomalies referenced to a specified reference surface, both at the earth's surface and in space.

B. The height of the geoid above a specified reference surface.

C. Deflections of the vertical at the earth's surface and in space.

In addition, requests for different types of gravity anomalies are anticipated — for example, isostatic anomalies. Thus, models for the elevation of the earth's topography and for the depth of compensation are required. Information on mean height above sea level is currently available for the entire surface of the earth, either as $5° \times 5°$ area means or as a spherical-harmonic expansion up to degree 35.

Some collection of surface-gravity measurements should be tabulated, including the methods of reduction and averaging that have been used. The data will most likely be in the form of area means, possibly from $1° \times 1°$ squares to $5° \times 5°$ squares. Data programs should exist for converting the observations from one reference surface to another and for converting different types of gravity anomalies.

There are already some determinations of time dependence of the spherical-harmonic coefficients in the geopotential. These should be included, when desired, in the potential representation.

2.7 Geology and Global Tectonics

As space and related techniques yield more and more detailed information about the earth, the geological value of the data increases correspondingly.

Table 5. Information for the earth's gravitational potential.

Service	Description	Size (characters)	Classification
Spherical-harmonic coefficients	Provides several sets of currently used normalized coefficients of the spherical-harmonic expansion of the geopotential — e.g., those obtained from the 1969 Smithsonian Standard Earth (II) (Gaposchkin and Lambeck, 1970).	10^4	Base
Surface-gravity measurements	Provides one or more sets of surface-gravity data, probably as area means over $1° \times 1°$ or $5° \times 5°$ squares.	10^6	Base
Free-air gravity anomalies from the geopotential model	Computes gravity vector at specified coordinates, using standard representation.	10^5	Process
Geoid profile	Computes geoid height above a specified reference surface, using standard representation.	10^5	Process
Tidal geopotential variations	Represents tidal dependence of the geopotential.	10^5	Base
Annual geopotential variations	Represents annual, and eventually other, long-period variations of the geopotential.	10^6	Base

Simultaneously, geology is in a state of transition. Once the hobby of Victorian clergymen, geology is now increasingly becoming a science of mathematical precision. There will always be a need for field geology, the source of the science's raw data. Nonetheless, the applications of computer processing to many aspects of geological analysis are becoming notable (Table 6), and there are some promising fields where such application is still awaited.

Two fields of computer application in geology are already well established. In the first, Bullard, Everett, and Smith (1965) have used a least-squares method to reconstruct the original proximity of the continents before the opening of the Atlantic Ocean. The success of the method, confirmed by Smith and Hallam (1970) for a southern ocean continental reconstruction, is a tribute to computer-assisted approaches to problems relevant to geology.

Second, the increasing number of published chemical analyses of rocks has aroused interest in the question of whether and how volcanic chemistry can vary according to tectonic situation. Chayes (1963) established a distinction between intra-oceanic and circum-oceanic basalts on the basis of TiO_2 percentage. Since then, he has refined his statistical methods, and he and other authors have shown how powerful a tool such discriminants can be in volcano-tectonics, particularly in the use of combinations of several variables plotted on n-space models (Chayes, 1969; LeMaitre, 1968). Several unexpected and most interesting results have already emerged from these approaches to chemical volcanology in space and time (e.g., Mohr, 1970).

Computer techniques have also found application in the field of paleontology, both in statistical ecological studies and in defining extinct species on the basis of skeletal parameters. In the fields of crystallography, mineralogy, and petrology, there are numerous obvious routine applications of computing — for example, in calculating cell-size parameters from X-ray diffraction data, in plotting variable chemical parameters of minerals or rocks on ternary or quaternary diagrams, and in computing rock norms (Chayes, 1964). These, in turn, can lead to further investigation into possible interrelationships among elements during the generation or crystallization of magmas, in a way

Table 6. Information for geology and global tectonics.

Service	Description	Size (characters)	Classification
Plate tectonics	Provides 1° coordinates of plate boundaries and relative motions at these coordinates (brought up to date); 1° coordinates of transform faults and fracture zones; gradient and shape of overridden plates, on a 1/2° grid. A crustal-motion model using these constants is being developed at SAO now.	10^6	Base
Seismology	1. Lists data on earthquakes — date, origin, time, epicenter, depth, magnitude, plate boundary, and wave analysis, including travel-time errors for all observing stations. 2. Provides data on surface deformation associated with all earthquakes (where developed); length and orientation of fissure and fractures, if more than one orientation; relative strength of development, displacement along fractures both vertically and horizontally with sense of movement and dip of fault plane; any associated rollofolds, with orientation and amplitude.	10^7	Base
Regional tectonics	1. Provides geodetic, geological, and seismic data for continental fault zones — e.g., San Andreas. 2. Gives rate of vertical movement for any 1° square on the globe. 3. Gives basement structural trend for 1° squares on continents.	10^7	Base
Deep crustal seismic-refraction studies	1. Provides P-wave velocities and densities for all layers (with thicknesses) down to and including the upper mantle for 1° squares. 2. Gives depth to M-discontinuity, 1° squares.	10^7	Base

Table 6. (Continued)

Service	Description	Size (characters)	Classification
Stratigraphy	Provides thickness (current and estimated original) of sediments (marine and non-marine) and lavas (basalt, andesite, trachyte, rhyolite) of each geological period for 1/2° squares; direction of derivation and lithology of sediments; depth to basement.	10^7	Base
Volcanism	Provides information on active, dormant, and extinct volcanoes: name and coordinates, plate boundary (if any), years of eruption, type of eruption, products (physical and chemical) of eruption and their volume and distribution, associated seismicity, air waves, atmospheric contamination, gas analysis, etc.	10^5	Base
Geochemistry and petrology	Provides data on all igneous rock analyses by 1) rock types, 2) rock ages, 3) geographic region, 4) tectonic settings; petrochemical parameters. Lists rock types and where found (1/2° squares over all the globe), associations of differentiates, partial melting products, accumulates, assimilates, etc.	10^7	Base
Paleoclimatology	Gives latitude and longitude of present 1° coordinates during each geological period; geological evidence of climate, wind direction, glacial flow direction, etc.	10^6	Base
Paleomagnetism	Lists paleomagnetic pole positions for the various geological periods and epochs; relation of drifting continents to these poles (see Paleoclimatology); all paleomagnetic	10^7	Base*

Table 6. (Continued)

Service	Description	Size (characters)	Classification
Paleomagnetism (continued)	observations, locality, rock type, rock age (stratigraphic and radiometric); relevant properties of analyzed specimen: oxidation state of rock and of iron ores, susceptibility, saturation magnetization, Curie point and total heating curve, Q factor, stability factor, remnant magnetization, etc.		
Seismic-interval process	Determines the time interval since an earthquake event of comparable magnitude at a given location, allowing for the investigation of possible regular migrations of seismic (strain-release) fronts on a continental scale. The input data will consist of earthquake dates, origin times, epicentral coordinates, and magnitudes.	10^5	Process
Plate-motion process	Provides the ability to link any two given points on the earth's surface in terms of the results of crustal-plate motions. This program will utilize the plate-tectonics data base, which will require an annual revision.	10^5	Process

that, once the essential picture of magmatic behavior is established, will enable detection of additional processes such as assimilation of the country-rock walling the magma chamber and thus contaminating the chemistry of the magma. Such hidden processes are at present very poorly known in even qualitative terms.

Yet another developing field of computer applications to geology is that of mechanical analysis of sediments, which can relate size- and shape-distribution patterns of the individual clasts (particles) to the source of sediments and can give information concerning their manner of removal and transportation and the conditions under which they were deposited. This aspect of geology is of obvious importance in oil exploration.

In geophysics, the applications of computers to gravity, seismic, and magnetic data reduction are too numerous and too obvious to list here, but the relating of factors such as the geoid to heat-flow or tectonic data, by means of harmonic analysis, is especially noteworthy in view of geology's new-found global vision.

The revolution wrought by the sea-floor-spreading hypothesis and the establishment of plate tectonics on a worldwide basis of crustal interaction have opened new fields for computer processing. Matching of continental margins is already well in progress. But the relating of volcanic mineralogy and chemistry to tectonic environment (in terms of plate tectonics) has hardly been started in global terms. There is an enormous wealth of suitable chemical data now scattered through the literature, and it would be a fascinating task to collect these, first to get a broad picture of global tectonic features and second to search for an explanation of "local" anomalies within this broad picture. This would almost certainly result in the discovery of features valuable to economic geology — for example, relating the chemistry, age, and tectonics of both diamond-bearing and nondiamond-bearing kimberlites.

A more distant but fascinating application of computers will occur when numerous precise geodetic data are available concerning the measured rates and sense of movements along major fault zones of the earth's crust. Already, the pattern of movement of the San Andreas fault zone has proved to be surprisingly complex (Hofmann, 1968), and the degree and manner of interaction of one section of this zone with adjacent sections is not yet understood. At a time when most or all of the earth's tectonic zones are being monitored, either on the ground or from ground-sensing satellites, a detailed analysis of crustal movement on the basis of global crustal-plate interaction, together with such factors as earth tides, isostatic loading, and heat flow, may possibly reveal a sequence of earth behavior that will enable the precise prediction of earthquakes (Mohr, 1969).

In summary, the possibilities of applications of computing processes in geology are almost as numerous as the specilizations of geology themselves. Where problems will arise is in the tedious collecting of published data, scattered through many journals (some of which are relatively inaccessible), and in the interpretation of processed data related to the fundamental, all-important field observations that will always remain the basis of a natural science like geology.

2.8 The Atmosphere

The earth's atmosphere enters an earth-physics program in several ways. In a very fundamental way, it enters because it is an important subsystem in the mechanics of the earth. Changes in the global distribution of atmospheric mass may account for some of the annual variation in J_2, the coefficient of the zonal harmonic representing the oblateness of the earth (Kozai, 1970). The annual variation in J_2, in turn, can account for perhaps half the annual variation in the rotation rate of the earth. The rest of the annual change in the rotation rate may involve momentum exchange between the atmosphere and the solid earth.

The upper atmosphere is responsible for significant aerodynamic drag on all satellites with perigees below perhaps 1000 km. Thus, knowledge of atmospheric density and its many changes at satellite altitudes is necessary for satellite orbital analyses.

The atmosphere is also a medium through which ground-based satellite-tracking observations must be made. Knowledge of the corrections for propagation of electromagnetic radiation through the atmosphere is necessary for the proper utilization of tracking data.

These several requirements indicate that a model of the earth's atmosphere is an important constituent of an earth-physics program (see Table 7). In the following, we shall discuss some of the existing and planned reference models.

The U.S. Standard Atmosphere 1962 is a detailed tabulation of atmospheric properties for a single standard profile from sea level to a height of 700 km. Since the molecular-temperature profile as a function of geopotential height is made to consist of straight-line segments, all quantities can be derived by straight algebra, without recourse to numerical integration. The tables are quite good to about 90 km but are obsolete above that height.

A series of realistic tabulations for different atmospheric conditions, taking into account seasonal and latitudinal variations in the homosphere and solar, geomagnetic, and other effects (based on Jacchia's 1965 models) in the heterosphere, is given in the U.S. Standard Atmosphere Supplements, 1966. From sea level to a height of 120 km, segmented molecular-temperature profiles are used in the same manner as in the U.S. Standard Atmosphere 1962 model; above 120 km, the temperature profiles are exponential functions of geometric height and a numerical-integration program was used to derive densities. The models enable a computation of atmospheric properties in the thermosphere and exosphere to be made in a continuous manner for all latitudes and local times.

A tabulation of atmospheric properties from 30 to 800 km appears in the COSPAR International Reference Atmosphere 1965 (CIRA, 1965). Only one average profile is given from 30 to 120 km, in segmented form. Above this height, detailed tables are given at intervals of 2 hours in local time for 10 separate levels of solar activity. In spite of all the detail, the tables are

Table 7. Information for the atmosphere.

Service	Description	Size (characters)	Classification
U.S. Standard Atmosphere 1962	Provides detailed tabulation of atmospheric properties for a single standard profile from sea level to 700 km. The model is quite valid up to 90 km but obsolete above this height.	10^5	Base
U.S. Standard Atmosphere Supplements, 1966	Based on U.S. Standard Atmosphere 1962, provides for different atmospheric conditions; seasonal and latitudinal variations in the homosphere; solar and geomagnetic effects. The models provide for computation of properties in the thermosphere and exosphere continuously for all latitudes and local times.	10^5	Base
COSPAR International Reference Atmosphere 1970	Gives the most up-to-date model of atmospheric properties in height, divided into three sections: 20 to 60 km, 60 to 120 km, and 90 to 2500 km. This models is ideal for satellite work but requires numerical-integration computer programs above 90 km.	10^6	Base
Atmospheric-model generator	Consists of a numerical-integration program to produce atmospheric density and composition as a function of position and time by means of input from the CIRA 1970 tables.	10^6	Process
Tropospheric correction	Allows angles, ranges, and range rates to be corrected owing to the refraction of electromagnetic radiation through the troposphere. Input parameters include observed quantities and geographical and meteorological data.	10^5	Base
Atmospheric mass	Provides data on the contribution of the mass of the atmosphere to the geopotential in space and formulas for correcting the sea-level equipotential for the mass of the atmosphere.	10^5	Process

valid only for one intermediate latitude and cannot be made continuous in latitude for satellite work. The computation is based on the integration of complicated equations and must be handled by numerical integration.

A revision of the COSPAR atmosphere is nearly ready and should be published in 1971; it should eliminate most of the drawbacks of the 1965 models and represent a considerable improvement over other existing models. At present, it is divided into three sections — 20 to 60 km, 60 to 120 km, and 90 to 2500 km; some interface adjusting has yet to be done to make the models continuous in height. Seasonal and latitudinal variations are taken into account in the homosphere and in the lower thermosphere; in the heterosphere, static models are used, with empirical temperature profiles that make numerical integration necessary in order to be able to derive data on density and composition. Computer programs to generate models above 90 km are available at SAO; for lower heights, the computation is trivial.

2.9 Oceanography

The oceans have a significant influence on the mechanics of the earth, and hence certain aspects of oceanography are pertinent to an EPIMS.

Historically, the mean sea surface has defined a reference from which to measure elevations on earth — i.e., elevations above sea level (see Table 8). Even with more sophisticated techniques such as satellite-to-ocean altimeters, the ocean surface can still provide a useful reference in this respect.

The ocean surface seems to be an equipotential surface to within a meter or so, and measurements of its profile by satellite altimeters can provide a significant improvement in the detail of a standard geopotential representation. Where available from surface measurements, astrogeodetic profiles of the ocean surface can be compared and combined with profiles obtained from satellite observations. For example, the sea over the Puerto Rico Trench is known to have about a 15-m depression relative to a reference ellipsoid of revolution (von Arx, 1967).

Table 8. Information for oceanography.

Service	Description	Size (characters)	Classification
Sea level	Gives data on the equipotential surface that best approximates mean sea level, as derived from the geopotential models, calculated as a function of latitude and longitude.	10^6	Process
Astrogeodetic profiles	Provides available astrogeodetic profiles of the oceans from surface measurements.	10^7	Base
Ocean trenches and other bottom features	Catalogs characteristics of ocean bottom features, such as ocean trenches, expected to have a significant signature in the geopotential.	10^5	Base
Hydrographic corrections to sea level	Presents corrections for deviations between sea level and geoid due to water temperature and salinity.	10^5	Process
Ocean currents and general circulation	Describes characteristics of major ocean currents. Provides differences between isobaric and equipotential surfaces.	10^6	Process
Tides	Provides a model of tidal variations of the sea profile.	10^6	Process
Tidal loading	Gives amount of tidal loading of crust (for use in analysis of body tides).	10^6	Process

Corrections of a maximum of a few meters for deviations between sea level and the geoid due to water temperature and salinity can be estimated (Hela and Lisitzin, 1967), and a model of these should be available in the EPIMS. Similarly, major ocean currents may produce a deviation of a meter or so between sea level and an equipotential surface. For this and other purposes, knowledge of the characteristics of the major ocean currents would be useful in the EPIMS.

Tidal variations of the mean sea surface will be measurable by satellite altimeters, particularly in shallow areas near shore lines.

2.10 The Ionosphere

The earth's ionosphere forms an integral part of any earth-physics program, primarily because of its effects on the propagation of electromagnetic radiation through it. A large number of measurements in earth physics already utilize satellite signals that have traversed the ionosphere (Table 9). It has become clear that some of the recently developed techniques, such as very long-baseline interferometry and laser ranging to satellites, will play a significant role in high-accuracy measurements in the next decade. Ionospheric effects on such measurements cannot be ignored, and a study of these effects (entailing a study of the ionosphere itself) is, therefore, of importance to earth physics. Since the ionosphere has a major variation with time, more or less continuous (hourly) data are necessary.

Although a study of the ionosphere per se would involve a very large number of parameters, the effect of the ionosphere on radio waves involves only those parameters that determine the refractive index n of the ionosphere. The refractive index differs from unity because of the presence of charged particles, the earth's magnetic field, and the collisional frequency of the electrons with neutral particles. Refractive-index variations cause both a bending of the rays and a change of the phase velocity of the signal. The bending effect is usually negligible, and the ionospheric effect can be expressed in terms of the differential phase path Δp (the difference between the path in free space and the actual path through the ionosphere):

Table 9. Information for the ionosphere.

Service	Description	Size (characters)	Classification
Electron-density measurements	Provides electron-density measurements for sites and times corresponding to pertinent earth-physics measurements through the ionosphere.	10^6	Base
Electron-density model	Provides the best model or models of the expected electron density as a function of position, altitude, time, etc.	10^5	Process
Refractive-index correction	Provides the necessary corrections for electromagnetic-radiation propagation through the ionosphere, based on either measured or modeled electron-density profiles.	10^5	Process
Integrated electron density along prescribed path	Provides the integrals $\int_0^S n(p)\,dp$ and $\int p\,n(p)\,dp$ necessary for refraction models for a given time and location.	10^5	Process

$$\Delta p = \int_{path} (1 - n)\, ds \quad .$$

For a rough calculation, the effects of the magnetic field and the collisional frequency can be ignored, and the refractive index can be represented by

$$n = \sqrt{1 - 81\, N/f^2} \quad ,$$

where N is the electron density (electrons/m^3) and f is the frequency of the radio wave (Hz). The strong frequency dependence of the ionospheric refractive index makes it possible to minimize ionospheric effects by carrying out measurements at relatively high frequencies (in the C and X bands). In such cases, the differential phase path is approximately

$$\Delta p \cong \frac{40.5}{f^2} \int N\, ds \quad \text{(meters)} \quad .$$

Thus, we need know only the total electron content of the ionosphere to be able to calculate the correction. It is also possible to use analytical models of electron-density profiles in the calculation of the differential phase path. However, true electron-density profiles differ widely from analytical models, and the latter can be used only for order-of-magnitude estimates.

For an earth-physics program requiring centimeter accuracy in range measurements, detailed knowledge of the ionospheric parameters is absolutely essential. Computer programs have been developed that calculate the differential phase path, taking into account the full anisotropic nature of the ionosphere. These programs use three-dimensional profiles of the electron density, the magnetic field, and the collisional frequency to calculate the correction for any given geometry of transmitters and receivers.

Although considerable data exist on the electron density of the ionosphere, there is an urgent need of a data-management program to coordinate the results of measurements made with a variety of techniques and in different geographical locations. The ionosphere is a constantly and rapidly changing region, and its parameters vary not only with height but also temporally and geographically. There is a need to combine measurements made on the upper and lower ionosphere by such techniques as topside sounding by satellites and ionosonde and incoherent backscatter radar measurements on the ground. Only when all these data are combined and organized in terms of geographical location and diurnal, seasonal, and yearly variations will it be possible to develop mathematical models of ionospheric parameters (the electron density in particular) that could closely predict the true state of the ionosphere.

At present, ionospheric data are compiled and stored by various agencies, such as the World Data Center in Boulder, Colorado. The basic sounding data are reduced to such quantities as, for example, critical frequencies of various layers and electron-density profiles; they are stored as a function of time and station location. No large-scale effort has yet been made to combine upper and lower ionospheric data to obtain comprehensive profiles. A systematic data-management program is needed to achieve this objective. When the ionospheric parameters are restricted to those having direct bearing on earth-physics measurements, the data-storage and data-handling requirements are more modest than might be appropriate for ionospheric research.

2.11 The Magnetic Field

A convenient representation of the earth's magnetic field is necessary for many earth-physics programs.

One of the most useful representations is a spherical-harmonic expansion, similar to the gravity-field representation (see Table 10). A data-management system should provide the user with a choice of several of the newest solutions for harmonic coefficients (solutions, possibly based on different types of data, may have different applications), as well as a simple program for evaluating the field at specified positions.

In cases where a dipole approximation to the magnetic field is sufficient, the dipole parameters (corresponding to the lowest order terms of each of the full solutions) can easily be made available.

Table 10. Information for the magnetic field.

Service	Description	Size (characters)	Classification
Harmonic coefficients for magnetic field	Gives several sets of coefficients for the representation of the magnetic field of the earth in spherical harmonics.	10^5	Base
Dipole approximation	Represents in the most simple form the geomagnetic field as a function of position and time.	10^6	Process
Detailed magnetic field	Represents the geomagnetic field as a function of all the harmonic coefficients, position, and time.	10^6	Process
Field-reversal tabulation	Lists the times of field reversals recognized from geological evidence.	10^4	Base

Time dependence of the magnetic field can be handled quite conveniently in the data system. Usually, in a spherical-harmonic expansion, the coefficients will be expressed as polynomial functions of time. It would be quite straightforward for the system to include routines for evaluating (and comparing) magnetic-field parameters as a function of time. It is likely that certain users will find surface magnetic-field measurements of interest, and these data, perhaps in smoothed form and with time-dependence parameters, should be made accessible.

Finally, another candidate for a data-management system is a listing of known field reversals of terra-magnetism (over geological time). These data (some 70 to 80 reliable reversals, and the number is growing) will be of interest not only to geophysicists interested in magnetism but also to those concerned with sea-floor spreading and correlations with other earth-physics data.

2.12 Ground Instruments

It would be desirable for EPIMS to provide a directory of stations, listing the characteristics, coordinates, and instrumentation of each. Something on the order of the NASA Geodetic Satellites Observation Station Directory is envisaged, but the EPIMS Directory should be broader, including polar-motion-determination sites, UT1-measurement sites, seismic stations, and other similar stations (see Table 11).

Two types of instrument coordinates will be of primary importance: 1) geocentric Cartesian coordinates and 2) geodetic datum coordinates (latitude, longitude, height above the geoid, and height of the geoid above the adopted reference ellipsoid). In addition, definitions of the coordinate systems are required — e.g., the scale, orientation, and dimensions of the reference ellipsoid. Accuracy information should also be available.

The basic data can be requested in various alternate forms, for example, as ellipsoid coordinates of an earth-centered ellipsoid of specified parameters

Table 11. Information about ground instruments.

Service	Description	Size (characters)	Classification
Directory of earth-physics instrumentation	Provides data on past, present, and future instruments and locations. The directory will facilitate scientific planning for such efforts as an international worldwide laser satellite-tracking network.	10^6	Base

or as Cartesian geodetic coordinates. The relationship between the geodetic and geocentric reference systems should also be made available, both as actual parameters and as the definition of the transformation. Facilities should then exist for converting any set of geodetic coordinates into a geocentric system once the transformation parameters are known.

Astrogeoid-height information should also be available in the data-management bank since such information forms a basic requirement for coordinate conversion. Currently, these data are available in the form of maps, but it may be useful to store the data in the computer as area means of, say, $1/2° \times 1/2°$ or $1° \times 1°$ squares. Astrogeodetic-height information is available for about 15% of the earth's surface.

2.13 Spacecraft

At any given time, many spacecraft that can contribute to an earth-physics program are in orbit. Data on their orbits and other characteristics should be available through EPIMS (see Table 12).

For satellites that have been observed previously, knowledge of the existing data and where it can be found would be desirable. It would not seem to be appropriate to expect EPIMS to serve as a data bank.

Table 12. Information on spacecraft.

Service	Description	Size (characters)	Classification
Directory of satellites in orbit	Provides current orbital parameters for prediction calculations, visibility patterns, and other processes used in scientific planning and analysis. The directory also provides information on spacecraft parameters, e.g., configuration sensors, shape, stabilization frequencies.	10^6	Base
Availability of satellite-tracking data	Provides information on what tracking data exist for which satellites and where they can be found.	10^6	Base
General prediction generator	Provides a set of programs to compute predictions on satellites using data on location coordinates, time, and observing instrument.	10^5	Process
Differential orbit program	Allows observed quantities to be compared against computed values based on satellite orbits.	10^5	Process

3. EARTH-PHYSICS INFORMATION-MANAGEMENT SERVICE

From Section 2, we can see that the information bases and processes available today in the earth-physics disciplines are numerous, complex, dynamic, and large. The potential users of the information are as diverse in character as are the disciplines themselves. In this section, we shall consider the nature of and problems involved in forming an information service that can match the wealth of information available with prospective users and provide a facility not currently available to the researcher in earth physics.

3.1 Concept of the Center

Over the past decade, two classes of facilities have come into prominence in the field of information management: 1) the computer service bureau, and 2) the information retrieval center or special library. Both structures have been greatly assisted by the development of the multiaccess, multiprogram computer system. If we consider a generalized information-handling system consisting of the primary sequential functions indicated in Table 13 (from Harrison, Brockman, and Sommers, 1970), the service bureau has in general assisted the researcher in achieving stages 2 and 3, and the special library, stages 4 and 5. A look at two NASA large-scale data systems further elaborates the delineation of current data functions and the implications to the individual researcher.

Table 13. Data-handling operations.

Stage	Operation
1. Acquire	Screen, exclude, select
2. "Process"	Condense, compress Digest, abstract, enhance (make more "visible")
3. Transform	Interpret, analyze, infer, deduce, project
4. Conserve	Update, index, taxonomically structure
5. Disseminate	Recall, search, match; "keyed" retrieval

The Information Processing Division (IPD) at GSFC receives and processes nearly all NASA's unmanned scientific satellite data. The IPD is virtually a totally centralized pure processing facility. It is primarily concerned with processing the data for scientific investigators, who in turn provide the necessary analysis. NSSDC complements IPD and is responsible for active collection, organization, storage, announcement, retrieval, dissemination, and exchange of space-science data. Figure 1 (from Harrison, Brockman, and Sommers, 1970) shows the relationship between IPD and NSSDC; whereas this flow of information is of value to research scientists, in general it suffers from the lack of the interactiveness to a broad community of users provided by modern computer technology and the lack of ability to provide a combination of information searching and subsequent processing in terms of an individual scientist's research.

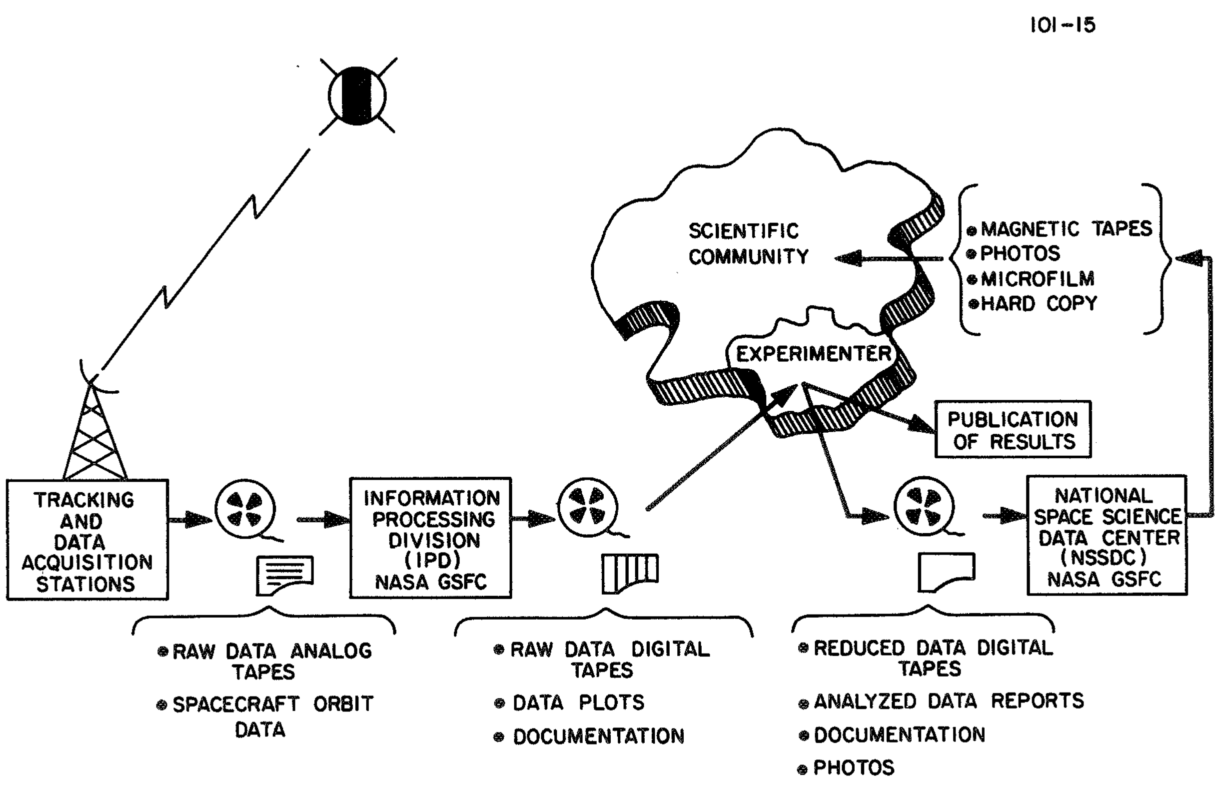

Figure 1. Space-science satellite-data flow.

If we consider the schematic information cycle for earth-physics research represented in Figure 2 (Lundquist, 1971) as the structure for EPIMS, then it is necessary to make the current models and algorithms and certain observations more readily available to the scientific community.

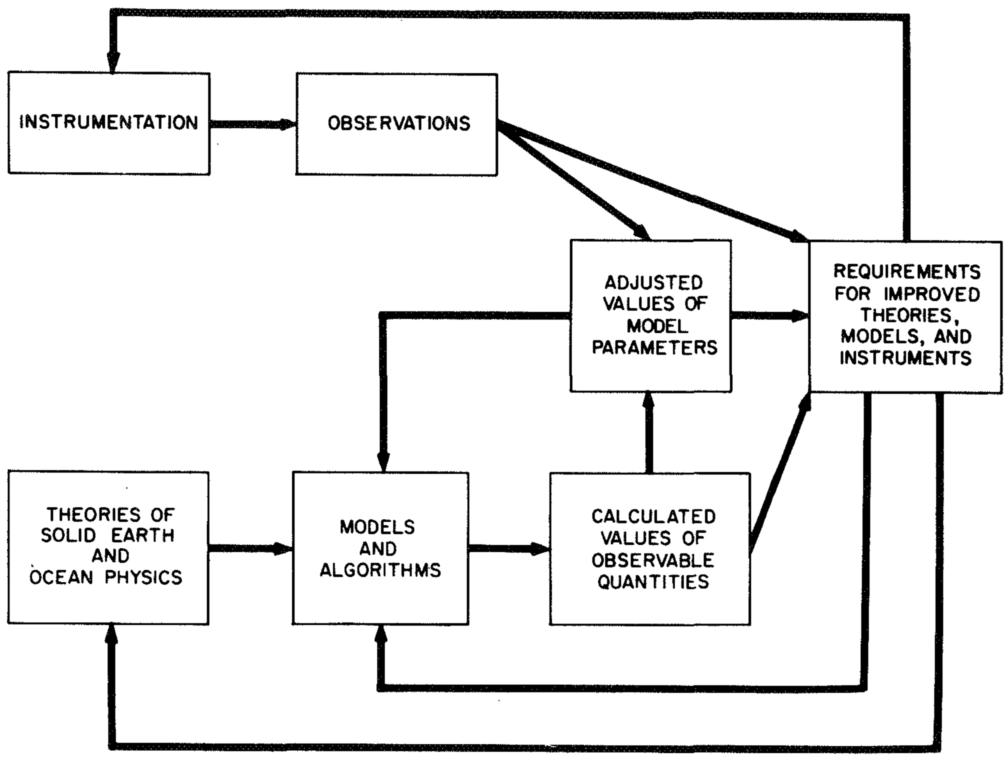

Figure 2. Schematic information cycle for application of space techniques to solid-earth and ocean physics.

Let us consider the case of a scientist who wishes to work with a model of the atmosphere as part of his earth-physics research. Depending on the size of his organization, he might have a good library or suitable computing facilities at his disposal. Or, through the efforts of a center such as NSSDC, he may know of results published by NASA. If he is successful in locating a suitable model, then with additional time and effort, he could perhaps obtain a computer-program version through yet another NASA facility such as the Computer Software and Management Information Center (COSMIC). With still a further expenditure of time, money, and effort, he could implement that program on the facilities at his own location or perhaps through a commercial service bureau. Hence, a typical scientist might, through considerable good fortune and with a minimum of three or four organizational contacts, achieve his original aims. The total inefficiency this process contains is compounded by the number of researchers and the organizations dealing in earth physics.

By combining the features of a computer service bureau and a special library through a centralized facility, EPIMS could achieve better results with considerably less effort and cost. Such a facility would provide a focus for both the active and the transient user in earth physics, as well as an interface service to the existing complex, heterogeneous, and sometimes confusing sources of information. All five information functions previously referred to would now be accessible in varying degrees to the general user. Figure 3 shows a conceptual structure of the center and its interfaces. In this structure, the three previously mentioned NASA facilities would become information sources for EPIMS.

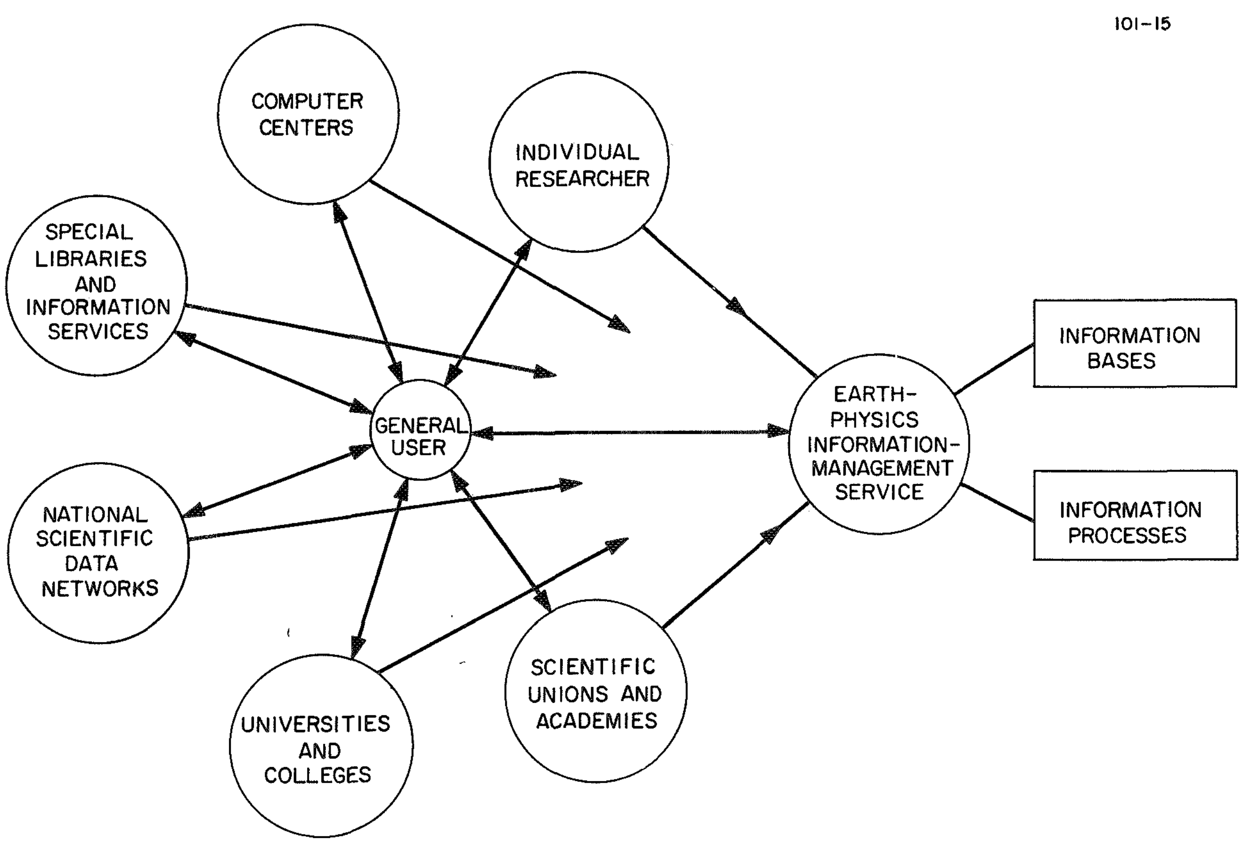

Figure 3. Conceptual structure of EPIMS.

EPIMS should provide a suitable approach to dealing with the concerns indicated by the Earth Surveys Planning Panel: "there are two aspects to consider: the job that must be done to assure that useful information generated as a result of (earth physics) program activities flows easily to the appropriate beneficiaries; and ... the extent (earth physics) information will be used by the full range of prospective users — governmental, institutional and private" (Tepper, 1969).

The rest of this section will amplify the nature of EPIMS, discussing, in turn, user access, information storage and availability, communications, the central computer system, protection of information, and economic considerations.

3.2 User Access

We anticipate that EPIMS will become a general contact point for the scientific community for the coordination of earth-physics information. As such, the center would interface with individuals and organizations having a wide range of capabilities and needs. In general, we could classify the access modes into three types: 1) no terminal, 2) interactive terminal, and 3) multi-device terminal.

3.2.1 No terminal

The users who do not require or do not have a terminal to the center would largely be researchers at random not directly involved in earth physics or those whose volume is so low as not to warrant a minimal terminal. In general, we expect that their requests would be received by phone or mail, and the center must have adequate staffing and query terminals so that these requests can be processed in the same manner as would requests from a user who has his own console. From the Smithsonian's experience with the management of the Scientific Information Exchange, the Center for Short-Lived Phenomena, and the Central Bureau for Astronomical Telegrams, individual information requests may be greater than 100 per day. This number is not unreasonable since, to provide maximum benefits, the center should be available not only to national researchers but also to the entire international scientific community.

Each request would obviously require a different amount of both man and machine processing, but less lead time in general would be needed by the non-terminal user than by the on-line user. An implication of the processing of nonterminal requests is the smoothing of computer workload on the shifts when interactive computing is not being provided.

It will be necessary to publish a directory of services, data processes, and bases available, particularly to aid the non-terminal user. We wish to stress, however, that publication is not a goal of the center, except in this limited context.

3.2.2 Interactive access

Most users will employ EPIMS through one of a number of commercially available inquiry/response terminals, either of the teleprinter or cathode-ray-tube (CRT) type. We expect that a user would, in general, query the system through the console as to the availability of various information models, processes, and data bases. Following this information search, he would find either a suitable model available to him with appropriate format options for his use or directive information on other possible sources to search. User/computer interactions would continue until this process had been exhausted. If the user wished to elaborate on an existing model and not use a "package program," or to interface a particular segment of a data base with his own program, he could develop his program in one of the languages available, e.g., BASIC, FORTRAN, ALGOL, etc.

If the user requires a considerable volume of information that is printed, punched, microfilmed, or in other ways hard stored, either this information would be generated at the center and supplied to the user or he would have to have local access to a multidevice terminal.

To provide the widest possible use of terminal consoles, the center would have to be able to interface with most standard inquiry/response terminals available today. We hope that with time, the emphasis might center on CRT terminals, which provide both alphanumeric character displays and line presentations. CRTs are superior as interactive terminals, with storage, error-correction, editing, and control capabilities not found on hard-copy consoles such as teletypewriters. However, hard copy is valuable and mnemonic, and combined with the number of teletypewriters in existence, it will probably remain the basic terminal for some time. CRTs with limited hard-copy-printer attachments that are now available would provide the ideal combination.

Tables 14 and 15 (from Computer Industry Annual 1969-70) indicate the characteristics and costs of typical teletypewriter devices and CRT consoles. The cost per terminal per month is between $100 and $500, including controller, with data modems costing another $25 to $50. Purchase prices generally range between $1,500 and $15,000.

Recent studies (Allen, Gerstenfeld, and Gerstberger, 1968; Utterback, 1969) have indicated that technologists turn first to the most accessible information channel, independent of the expected value of the information it will provide. In particular, this is the case once a problem has been identified. Although the information search and use behavior between scientists and technologists are different, the general consensus is that accessibility is an overriding determinant in the selection of an information source. In other studies (O'Gara, 1968; Frohman, 1968) concerning the physical distance separating individuals and nonhuman communication points such as computer consoles, it was found that the probability of use decreased with the square of the distance. The probability was asymptotic within 25 yards. Such user/device accessibility relationships indicate the desire for consoles with characteristics of light weight, mobility, low cost, rapid communication tie-in, and simplicity of commands for system entry.

In conversational computing, we can define an interaction cycle between the user and the computer system as composed for four basic elements — think, input, response, and output. We can define the action time as being the sum of the output, the think, and the input elements. When a user is interacting

Table 14. Typical teletypewriter console characteristics.

Manufacturer	Computer Displays Inc.*	Control Data Corp.	IBM	Univac
Model	ARDS	200 series	2265/2845	Uniscope 300
CRT Characteristics				
Screen size	6.5" × 8.5"	6" × 8"	9" × 7.5"	10" × 5"
Phosphor	P1	P4	P39	P31
Spot diameter	0.008"	–	0.018"	0.012"–0.020"
Brightness (foot-lamberts)	3	75	–	50
Display Format				
Char/line	80	50 or 80	64 or 80	64
Lines/display	58	20 or 13	15 or 12	8 or 16
Max. char/display	4600	1040	960	1024
Character size and generation method	0.08" × 0.06" stroke	0.25" × 0.125" 5 × 7 dot matrix	0.178" × 0.126" stroke	0.15" × 0.113" stroke
Memory Buffer				
Type	Storage tube	Delay line	Delay line	Core
Size (characters)	4600	1000	960	1024
Refresh rate (frames/sec)	not required	50	54	60
Editing Features				
Insert	No	Character	No	Character or line
Delete	No	Character	Character	Character or line
Cursor	Yes	Yes	Yes	Yes
Special Controls	Nonstore erase, tabs	–	–	Split screen tabs
Transmission				
Data set interface	201, 202, 103, sync. or async.	201 A, B, C, D	201 B, 202 D	201 A, B (sync.)
Speed (bits/sec)	2400 bps	2400 bps	1200 or 2400 bps	2000/2400 bps
Simultaneity	Half or full duplex	Half duplex	Half duplex	Half duplex
Codes	ASCII	BCD/ASCII	ASCII	ASCII
Error detection	Parity	Yes	Yes	Parity
Alternate Outputs	Yes	–	Printer 1053	Printer
Special Features	Graphic capability; vector generator; mouse or joystick available	–	2845 is controller	40 function keys
Comments	Stand-alone terminal	Stand-alone or multistation	Stand-alone terminal	Stand-alone or multistation
Multiple Terminals	Yes	Up to 12	–	Up to 48
Cost				
Purchase price	$7,950	–	$5,600/display $8,550/control $2,000/printer	$15,140/single station $268,000/max. config.
Rental/mo.	None	$135/terminal $724/controller $270/printer	$175/display $175/control $50/printer	$350/single station $8,330/max. config.

*These are storage CRT terminals and have limited graphic capability.

Table 15. Typical CRT console characteristics.

Manufacturer	IBM	Teletype	Teletype
Model	1050	32/33	35/37
Input Mechanisms	Keyboard Paper tape reader Card reader	Keyboard Paper tape reader (optional)	Keyboard Paper tape reader
Output Mechanisms	Paper tape punch Edge punch cards Card punch	Page printer Paper tape punch (optional)	Page printer Paper tape punch
Formats	51/80 col. words	72 char/line width (max.)	72 char/line (max.)
Document Characteristics	8.5" wide paper	8.5" wide paper 1 paper tape	8.5" wide paper 1 paper tape
Codes			
I/O medium	7 level	5 level (Model 32) 8 level ASCII	Any 8 level
Transmission medium	7 level	5 level (Model 32) 8 level ASCII	Any 8 level
Transmission			
Speed	10–15 char/sec	6–10 char/sec	6–10 char/sec
Simultaneity	Half or full duplex	Half or full duplex	Half or full duplex
Synchronization	Asynchronous	Asynchronous	Asynchronous
Error Detection			
Lateral/Longitudinal	Lat/Long	None (Model 32) Lateral (even)	Lateral (even)
Corrective action	Automatic retransmission Delete Repunch	Manual retransmission	Manual retransmission
Comments	Friction feed, standard; Pin-feed RPQ	Pin-feed paper Also available: 12 char/inch printing	Pin-feed standard; Friction feed available
Cost			
Rental/mo.	\$100–\$500		
Purchase price		\$400–\$650	\$1,400–\$4,000

heavily with the system, it has been observed that the action time averages about 20 sec (Hyman, 1967) on a wide range of systems. It has further been observed that if the response time exceeds 10 sec, interaction becomes intolerable. A response time of up to 2 sec seems to provide the desired continuity for user/machine interaction. A typical response curve for a medium-scale drum and disk system is shown in Figure 4 (from Hyman, 1967). In general, the degree of responsiveness requires a consideration of the tradeoff between the relative use of the central memory and the secondary storage transfer rates. Characteristics of the central computer system are discussed in more detail in Section 3.5.

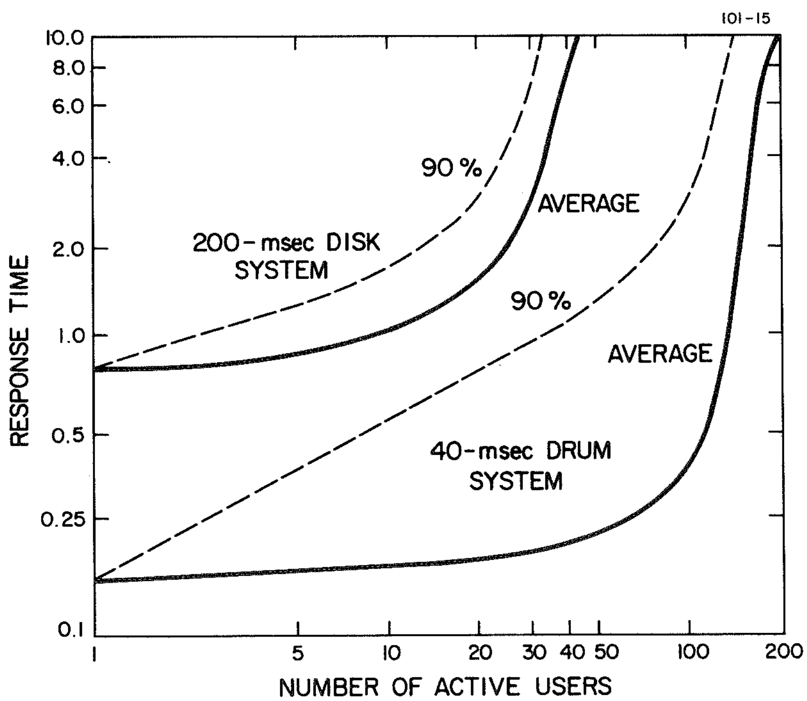

Figure 4. The interaction cycle.

It would seem reasonable, based on the size of the earth-physics research community, to consider that EPIMS should provide simultaneous console interfaces for 25 to 50 users initially, with the capability to increase this number to 100 or 200 during the early phases of operation.

3.2.3 Multidevice terminal

Some users of EPIMS will have access either to a computer system or to a high-speed terminal that can act as a remote station. In other cases, there may be a group of users geographically close that would warrant installation of terminal devices other than interactive consoles. In this context, a multidevice terminal would probably be made up of a line printer, a card punch, a card reader, and optional magnetic tape units. Data transmission would be carried on in the range of 2000 to 28,000 bits per second. We expect there would be a limited number of such user facilities. It would seem reasonable to have the capacity to handle five such devices initially.

The center would expect to utilize interfaces, such as those outlined, with other information-processing facilities where the volume of data to be exchanged or shared would warrant it.

In general, multidevice terminals require an environmentalized location with properly trained personnel to operate and supervise the equipment. Typical configurations would rent for $2,500 to $6,000 per month. Comparable purchase prices would range between $100,000 and $250,000.

3.3 Information Availability and Storage

Considerations of information storage go beyond the physical volume required. It will be necessary to develop techniques, procedures, and guidelines that consider such questions as who should determine what is stored and available and why, how much information is required, and for how long information should be available.

Of necessity, the personnel and staff of an information center such as EPIMS will become technological and scientific gatekeepers. This extremely difficult role will require a broad awareness of the state of scientific development in many fields. The task will be laced with the subtleties required in dealing with a heterogeneous mix of national and international organizations and individuals. If, for example, there are six currently accepted atmospheric models, but only two are considered necessary for the purpose of such an information center, the care required to handle such choices cannot be understated.

EPIMS will have to have on its staff or available to it a number of scientists who will assist in the information-selection process. Input must come largely from the user community, for it is only for them that the center would exist. Hence, requirements for and use of information processes or data bases would become part of an evolving process involving acceptance of and searches for information through scientific unions, individual researchers, universities, colleges, etc.

In general, it would seem that the location of such a center should be at a facility having an existing scientific staff working in a broad range of disciplines associated with earth physics. That facility should have existing library and data-processing and communications facilities. Scientific and technical management experience in dealing with complex data-processing projects from inception through completion would be highly desirable. An organization that is neutral regarding national scientific-policy setting and implementation would be advisable, in order to encourage international participation.

The volume of information that could be stored and made available through EPIMS is already considerable. Over the next decade, the possible growth factor is enormous. This is further complicated by the varying degrees of attentuation that apply to earth-physics information. Nonetheless, when vast quantities of information are being dealt with, certain guidelines must exist for purging the system and keeping the available data within manageable limits. Gurk and Minker (1970) have developed a stochastic model relevant to information-handling centers typified by computer utilities and document storage and retrieval such as we have been describing. The growth characteristics of such an information center are evaluated for a retirement policy that governs when items should be retired from a two-level auxiliary store (disk or drum, in the case of the computer utility) to a less accessible store (tape or microfilm). Retired information can be reactivated into the primary store, provided a sufficient number of requests have been made for it. Simply stated, the retirement policy says that an item in the primary store is retired if it arrived more than T years ago or if it has been in the primary store at least X years ($X \leq T$) and has been used less than K times in the previous Y years ($X \geq Y$). The rebirth policy would require that an item requested from the secondary or retired store be placed in the primary store if it has been requested at least K times in the previous Y years, provided it did not arrive more than T years ago.

Storage policies such as the above should be applied to the information bases available to the user as well as to the individual user storage he might require for his programs or personal file storage.

In certain cases, the question of providing storage can be viewed as a tradeoff between the cost of data transmission and that of storage facilities. We anticipate that, where practical, such a center should not duplicate data bases that might be accessed elsewhere. When a data base already exists on-line to a computer system at another center, an evaluation should be made as to whether a sufficient volume of use dictates establishing a telecommunications link between the two computer systems.

Judging from Section 2, the information in EPIMS would probably range between 10^{10} and 10^{11} bytes in its early stages and would be stored in various modes. The primary storage should consist of a combination of rotating mass memories, such as magnetic disks and drum devices. A suitable complement of magnetic tape units will also be required. A combination of these devices will provide the capability for efficient and economical manipulation of both random and sequential accessible information. Each technique exhibits its own characteristic speed of access, storage capacity, and cost per bit. For a center such as EPIMS, the principal emphasis will be on large on-line random-access storage, through rotating mass memories.

Mass memories are either of the head-per-track or positionable-head configuration. The average access time on fixed-head devices ranges between

10 and 30 msec with a storage capacity available up to approximately half a million bits. Average access time for positionable-head devices ranges between 50 and 500 msec with a storage capacity of greater than 5 billion bits. The cost for the larger fixed-head device is approximately $150,000, and up to $400,000 for the larger positionable-head units.

Standard 2400-ft magnetic tape reels at 800 BPI (bits per inch) would store approximately 20 million bytes of information. Magnetic tape would most likely be the basic secondary-storage medium. A third or perhaps alternate secondary storage could be accomplished with microfilm.

Other new classes of storage equipment are currently available or under development, ranging from magnetic data cells to laser-written memories. An information center such as EPIMS will constantly evaluate its ability to provide service to its users, taking the most economical and efficient advantage of new technology as soon as it has proved its reliability and performance.

Although in certain cases a smaller number of storage devices might hold the physical volume required, it will probably be necessary to trade off volume against the number of access channels to storage devices. Two mass memories with two access ports will provide considerably different queuing times than would four devices with four access channels and the same equivalent storage. Duplicate devices also provide for a better continuity of operation from either planned or unplanned down time.

An operational consideration involving mass storage is the ability to dump periodically all files that are being stored on large disk systems. In general, this dumping is done to magnetic tape every hour or so, providing a continuity of operation that would otherwise be lost owing to inadvertent operator error, machine malfunction, etc. It will be necessary, therefore, to provide such dumping capability parallel to normal operation so that the quality of service will not be degraded.

The case of off-site backup storage to prevent catastrophe is not of principal concern to a center such as EPIMS, since the processes, data bases, and information handled in the center would, in general, also be available elsewhere.

3.4 Communications

For any large-scale on-line data-processing center, communications costs can be considerable if the center derives a significant portion of its workload via remote terminals. Considering also the impact of communications upon system performance and future growth potential, it is clear that a significant amount of effort should be dedicated to the design of an optimum communications system.

Since telephone companies provide a variety of data-transmission services (or channels), one of the more obvious decisions that must be made by the designer is the proper choice of channel. In essence, this decision will be made in conjunction with the selection of the remote peripheral equipment, since the designer will wish to match the line capabilities with the speed of the peripheral device with which it is associated. Table 16 lists typical common carrier channel characteristics and Table 17, common carrier services (from Computer Industry Annual 1969-70). It should be noted that, currently, leased conditioned lines are necessary for speeds above about 2000 bits/sec to ensure proper quality of data transmission.

Since the propagation of D.C. signals is impractical except for very low-speed transmission rates over short distances, modems (modulating-demodulating equipment) are utilized to transform binary signals into the analog signals of the transmission lines and vice versa. Various types of modems and transmission modes must be considered. Modems operate in a parallel mode or in a bit-serial mode. Except for some low-speed terminals, almost all terminals operate in the bit-serial mode. Since computers and their associated I/O devices communicate with each other in parallel mode, communications interfaces must be provided with bit-serial interfaces to perform both serial-to-parallel conversion and parallel-to-serial conversion. Another convention that must be dealt with is the communications code. The eight-level USASCII code is highly recommended. While the Baudot five-level code is fairly common, especially in older equipment, USASCII has gained wide-spread use and has been adopted as a proposed standard, simplifying the intercommunication between equipment of various manufacturers.

Table 16. Common carrier transmission facilities.

COMMON CARRIER	FACILITY DESCRIPTION	BAND RATE (BIT/SEC.)	MODEM USED	NO. OF VOICE CHANNELS REQUIRED	COMMENTS
AT&T	Schedule 1 Teleprinter channel	45	None	1/12	Available on TWX
AT&T	Schedule 2 Teleprinter channel	57	None	1/12	Available on TWX
AT&T	Schedule 3 Teleprinter channel	75	None	1/12	Available on TWX
AT&T	Schedule 3A Teleprinter channel	150	None	1/8	Available on TWX
AT&T	Schedule 4 Voice channel	2000	Bell 201A	1	Available on dial network, various modems used
AT&T	Schedule 4A Voice channel	2400	Bell 201B	1	Various modems used
AT&T	Schedule 4B Voice channel	3600	Bell 203	2	Available on dial network
AT&T	Schedule 4C Voice channel	4800	Bell 203	2	Other modems may be used
AT&T	Schedule 4D Voice channel	7200	Bell 203	2	Other modems may be used
AT&T	Broadband channel	19,200	Bell 303A10	12	Used with Telpak
AT&T	Broadband channel	50,000	Bell 303A20	12	Available on Data-phone-50
AT&T	Broadband channel	230,400	Bell 303A30	60	Used with Telpak
WESTERN UNION	Class A Teleprinter channel	50	W.U. 1181-A	1/12	Available on Telex
WESTERN UNION	Class B Teleprinter channel	57	W.U. 1181-A	1/12	Available on Telex
WESTERN UNION	Class C Teleprinter channel	75	W.U. 1181-A	1/12	Available on Telex
WESTERN UNION	Class D Teleprinter channel	180	W.U. 1181-A	1/8	Available on Telex
WESTERN UNION	2KC Broadband data channel	600	W.U. 1601-A	1/2	Available on W.U. Broadband exchange
WESTERN UNION	Class E Broadband data channel	1200	W.U. 2121-A	1	
WESTERN UNION	8KC Broadband data channel	4800	---	2	
WESTERN UNION	16KC Broadband data channel	9600	---	4	
WESTERN UNION	48KC Broadband data channel	28,800	---	12	Available on Dial-pak

Table 17. Common-carrier service tariffs.

COMMON CARRIER	TYPE OF SERVICE	BANDWIDTH or NOMINAL SPEED	TARIFFS* (Dollars)	COMMENTS
AT&T BELL SYSTEM	TWX dial network	100 words per min.	$1.75 first three minutes 0.60 each additional min. for 2000 miles and over 0.20/min. for 0–50 miles	Charges vary with distance, three minute minimum charge
WESTERN UNION	TELEX dial network	66 words per min.	From $0.175 to 0.60/min. depending on areas. 40% discount on excess if charges exceed $87.50/mo.	No 3 minute min. charge, fractions of min. are propor. charged.
AT&T and WESTERN UNION	Private line services low-speed	60, 75 words/min.	$1.10/channel-mile/month for half-duplex $1.21 for full-duplex	Charges are telescopic, reduce to half after 250 miles and third after 1,000 miles
		100 words/min.	$1.21 for half duplex, 1.331 for full duplex	
		sub-voice 150-180 bps	1.375 for half duplex, 1.513 for full duplex	
AT&T BELL SYSTEM	Public telephone network, dial exchange	voiceband 3000 bps	$1.00 to $2.00/three min. 0.25 to 0.50 each add'l. min. for 2,000 mi. and over $0.30/three min. and 0.10 each add'l. min. for up to 30 mi.	charged by time and distance 3 minute min. charge for call.
	WATS unlimited service	voice network 3,000 bps	$2,300/month for anywhere in USA for unlimited time. $ 500/month for Area 1	cost depends on service areas and the state.
	WATS measured service	voice network 3,000 bps	$610/month for first 15 hours, $34 each add'l. hour for anywhere in USA	cost of service lower for indiv. areas, varies.
WESTERN UNION	BEX Broadband Exchange	2 kc/s	$0.15/minute to $0.65/min. 40% discount on excess of $3000	Rates depend on areas, charges are broken down to tenths of a minute
		4 kc/s	$0.20/minute to $0.75/min. 40% discount on excess of $400	
AT&T and WESTERN UNION	voice-grade leased lines	4KHz	$2.02 for half duplex $2.22 for full duplex (channel-mile/month) $10, $37.50, and $56. channel condition charges for schedule 4A, 4B, and 4C respectively	rates are telescopic drop by approx. 15% after 250 mi. and 25% after 500 miles.
AT&T	TELPAK A	48 kc/s	$15/mile/month	12 voice channels
	TELPAK B	96 kc/s	20/mile/month	24 voice channels
	TELPAK C	240 kc/s	25/mile/month	60 voice channels
	TELPAK D	1,000 kc/s	45/mile/month	240 voice channels

*These are approximate tariffs currently available and may change. Check with your local common carrier representative.

Transmission itself may be synchronous or asynchronous. Little need be stated here about this, except to say that with asynchronous transmission each character or control code is transmitted independently, thus making it suitable for devices where characters are generated at irregular intervals, e.g., teletypes. However, synchronous transmission makes more efficient use of the transmission line, and thus it is capable of higher rates of transmission.

Error detection and correction can be logically divided into the two obvious problem areas of detecting the error and then correcting the data at the receiving terminal. Many schemes exist for accomplishing this via combinations of hardware and software. Suffice it to say that essentially any degree of reliability of transmission can be achieved, but one must also be aware that a significant price must be paid for increased reliability, in terms of either hardware or software overhead or both.

There is one area associated with communications in which substantial savings can be achieved — that of multiplexing or concentrating. Simply speaking, this technique can reduce costs by enabling users to transmit multiple data streams over one telephone circuit. In general, two basic techniques are utilized to accomplish multiplexing — frequency division and time division. The frequency-division technique divides the bandwidth of the voice-grade circuit into a number of narrower channels; the time-division technique assigns alternating slots of time to the various channels.

Techniques have been developed to accommodate a wide variety of remote-user configurations. For instance, if a number of remote users are in the same telephone-exchange area, all their terminals could be fed into a multiplexer at the remote site and transmitted over a single leased line to another multiplexer at the central computer site. However, a single circuit may also serve several cities, with channels being dropped at each point along the way. This is known as a multidrop configuration. Other techniques for somewhat special circumstances are the high-speed intermix and the shared multiplexer at the central site.

3.5 The Central Computer

From the nature of the computer-oriented information-management system being considered here, it is possible to specify in a general manner the major features of the desired computer system. It is evident that the hardware selected should permit the implementation of an interactive multiprogramed system in a straightforward and optimum manner.

Multiprograming is desirable from an economic point of view since it allows dynamic sharing of the available resources among two or more programs. Specifically, it permits the overlap of computing and I/O, which should be a major asset for the type of system being considered. The reasons for specifying an interactive system are many. Interactive systems are usually associated with time-shared systems, mainly because time sharing has been emphasized by those seeking to provide interactive access to a computer. The concept of interactive computing and, consequently, time sharing derived its main impetus from the lack of satisfactory debugging capabilities in a batch-processing mode. Currently, the real benefit of interactive computing is the ability to compile, debug, and run programs in one continuous session at a console. This has had an unexpected impact on research. In fact, Corbato and Vyssotsky (1965) have stated, "The availability of the MAC system has not only changed the way problems are attacked, but also important research has been done that would not have been undertaken otherwise."

The hardware should be appropriately selected as an integral part of an interactive multiprogramed system and should be modular as far as possible. Future expansion should include the possibility of multiple processors as well as the upward expansion of the main memory. Some general features of the hardware might be the following:

1. One or more central processing units accessing a common main memory.

2. Greater than 100,000 cells of location-addressed main memory.

3. Data channels capable of being interfaced to a wide variety of peripheral devices.

4. A secondary memory consisting of perhaps 10 million cells with block access time of the order of milliseconds.

5. A tertiary memory with block-access time several orders of magnitude greater than the secondary memory and capable of holding all planned files.

6. Clocks and meters for measurements on the system.

Most large-scale systems today have software that either is supplied with the hardware or can be obtained from the manufacturer under a separate agreement. The software suitable for this specific application should have certain characteristics, such as the following:

1. Higher level language compilers.
2. Dynamic debugging capabilities.
3. Text-editing capabilities.
4. Ability to communicate with a variety of remote devices.
5. Extensive file-handling capabilities.
6. Tutorial routines to instruct users on the use of the system.
7. An efficient and sophisticated supervisor program.

In general, the first three items should be standard, usable as supplied by the main vendor. Internal development may be required on the fourth to make the package supplied by the manufacturer more general.

Item 5 may require the most attention since files represent the major reason for the system's existence. The file structure must be random-access and names should be identifiable symbolically, both externally and internally. The handling routine must be able to locate files in a minimum of time and update files readily so that they are easily found and extracted. The packing of information must be reasonably high, and most importantly, it must possess an efficient file-addressing scheme (i. e., the transformation from a symbolic file name into an address on some mass storage device).

Item 6 should be prepared by the implementers of the total system and should reflect simplicity throughout, as should the commands these routines describe.

To a large extent, the satisfactory performance of the overall system will depend upon the skill with which item 7, the supervisor program, has been implemented. In general, these routines will be supplied by the hardware manufacturer. The supervisor program must be capable of keeping information in relocatable form until execution time (rather than load time), so that it does not matter where in the core a segment is loaded; the supervisor must be able to handle multilevel interrupts, dynamic scheduling, and allocation of priorities. It must be able to handle queues that will grow and diminish in a probabilistic manner and to handle overloads while allocating resources in such a way that the system does not come to a halt. Finally, but most importantly, the supervisor program must be highly reliable, for everything depends on its efficient and faithful operation.

3.6 Protection of Information

Special problems concerning protection of information arise when a computerized information system services a large community of users, many of whom may be using the system simultaneously. There will be problems associated with the sharing of data and processes as well as problems arising because groups of users may be working on cooperative efforts. While it may be true that "privacy of information" may not be an issue here, since there should be no real conflict with respect to users' goals, protection is still highly desirable for several reasons:

1. Protection of processes limits the propagation of errors throughout the system while a user is debugging or utilizing what was thought to be a debugged process.

2. Protection of information is necessary to prevent inadvertent (or malicious) changes or updating by improper authorities. This is quite necessary for reliable operation and reproducibility of results when identical tasks are being performed.

3. Protection is necessary for allocation of resources and usage records in order to ensure proper charging to users.

In general, it is probably safe to say that for efficient operation of a large system, hardware must bear some of the responsibility for protection. One solution in the past has been to provide a mode switch that enables processes to run in a master-slave environment with certain instructions inaccessible to processes running in the slave mode. I/O instructions usually fall within this class. This mode switch has been utilized in conjunction with memory-bounds registers, which limit the address space in working core for a process. That is, these registers delineate a specific block of core locations outside of which the specific process (operating in slave mode) cannot access. This implementation has one distinct disadvantage: If a process has access to information, the access is unrestricted. It may read, write, or execute; there is no distinction among a variety of access privileges or capabilities.

Many modern machines have been designed to accommodate a variety of access switches on physical memory segments. A variety of access rights, such as read, write, or execute, or a combination of these, can be specified on a particular physical segment. However, if a user has access to a particular segment, he has the same access rights as any other user with access to the same segment. This is not satisfactory for the case of shared segments, where many users may want to use the data or process but only a limited number of users should have the ability to modify or update the information. This has led to the implementation on the most advanced machines of access rights based on logical segments, allowing different access rights to different users on the same physical segment. Thus, a user (or process being utilized by a user) can be assigned the proper access rights to the proper information segments for a specific task. A user who creates a process or data segment may also specify those access privileges that he is willing to grant other users.

One particular point should be emphasized. Our own experience has indicated that it is absolutely necessary that there be one person who is administratively and totally responsible for updating the data base. He must be informed of and verify all updates (to see that they have been implemented correctly and on schedule) and must speedily dispatch all relevant information.

Finally, since malfunctions must be anticipated, the facility must provide backup storage for all important files and segments. This should be implemented so that the system can be reinitialized regardless of the malfunction.

3.7 Economic Considerations

In the preceeding, we have outlined some of the considerations for building an information-processing system that will provide maximum usefulness to the earth-physics scientific community. One of the key themes that has been implicitly stressed and should now be made explicit is the desire for modularity. Figure 4 graphically displays this concept and what should be considered the four basic components of the system — terminals, multiplexers, processors, and storage. A large system built from relatively simple modules functioning in parallel is inherently more reliable than a large "kludge," and even operating at a degraded level, such a system can still give partial service. Modularity allows a system to be built up or cut back at any or all levels, as a result of changing requirements, costs, services, etc.

In Table 18 we have summarized the costs of a system at the 100-terminal level. It seems reasonable to consider a system capable of this capacity to determine the total feasibility and acceptability of such a center as EPIMS. This number seems valid when the total number of potential users is considered, a number that could easily range to three or four thousand. Virtually every university or college conducting a program in earth physics is a likely user, at a minimal cost of $1,000 to $2,000 a year for equipment. The range of operational costs, between $117,000 and $152,000 per month, is consistent with a summary of typical systems of different sizes presented in the Computer Industry Annual 1969-70 (see Table 19). These costs refer to an average monthly expense once a reasonable state of operation has been attained. We anticipate that such a major information center would require approximately a year to 18 months to develop into a 100-terminal system.

Table 18. Expected annual operating expenses for EPIMS at the 100-terminal level.

Equipment	Purchase Price	Monthly Rental Price
A. Central Computer — featuring multiprocessing, multiprogramming, large memory for multi-function scientific interactive (100 terminal) and batch-job mix	\$2,000,000 — 2,500,000	\$40,000 — 50,000
B. Storage System On-Line		
2 large positionable-head disks, multiaccess channels, approximately 1.5×10^9 bits storage	500,000 — 700,000	5,000 — 9,000
1 drum-unit, fast-access device, to handle I/O queue	150,000 — 300,000	3,000 — 5,000
5 or 6 tape drives	150,000 — 300,000	3,000 — 5,000
C. Peripheral Devices — card readers, printers, controllers, etc.	200,000 — 300,000	4,000 — 5,000
D. Communications Equipment — including multiplexers and controllers at the computer center	150,000 — 200,000	4,000 — 5,000
E. Terminals		
50* inquiry-response	75,000 — 350,000	5,000 — 17,500
2 multiplexers	250,000 — 275,000	7,000 — 8,000
F. Miscellaneous Equipment — tape cleaner, storage files, data-handling devices, etc.	50,000 — 60,000	1,000 — 1,500
Range:	\$3,525,000 — 4,985,000	\$72,000 — 107,000
Yearly Rental:		\$864,000 — 1,284,000

ADDITIONAL COSTS

Personnel Compensation			
	Manager	\$25,000	
5	Discipline Researchers @ \$17,000	85,000	
3	Applications Programers @ \$13,000	39,000	
3	Systems-Development Programers @ \$15,000	45,000	
2	Keypunch Operators	11,000	
2	Information Research Specialists	24,000	
1	Secretary	6,000	
4	Computer Operators (2 shifts)	40,000	
			\$285,000
Personnel Benefits @ 10%			28,000
Travel			12,000
Facilities — Space, Utilities, A/C, etc.			45,000
Supplies			30,000
Equipment Maintenance (average for equipment schedule above)			145,000
TOTAL COST/YEAR			\$1,409,000 — 1,829,000

*It is assumed that a number of users already have terminals and that only 50 terminals would be obtained through the center.

Table 19. Monthly system costs.

	20 lines	50 lines	100 lines
Computer center	$ 8,800	$22,000	$ 41,000
Communications facilities	5,900	29,000	46,000
Terminals	1,500	5,000	10,000
Development (36-month writeoff)	3,300	9,000	17,000
Operation and maintenance	5,500	11,000	22,000
Total	$25,000	$76,000	$136,000

Growth beyond this size must be evaluated in terms of user acceptance and economic considerations. It will be necessary to establish prices of the services offered by EPIMS, but the principal emphasis during the early stages of development should not be directed at achieving a self-supporting level at the expense of developing a worthwhile service.

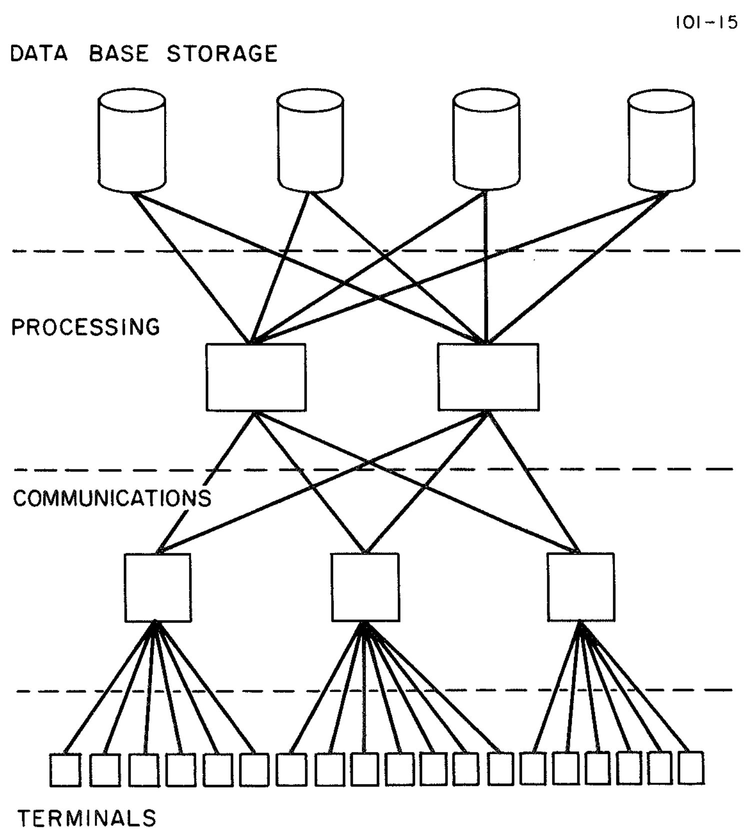

Figure 5. Typical system configuration.

4. CONCLUSION AND RECOMMENDATIONS

After reviewing the information bases and processes available to the earth-physics disciplines and the data-processing technology in existence today, we conclude that an earth-physics information-management service is both desirable and feasible. Further, we feel that the potential benefits to both the general and the specialized user of earth-physics information would be considerable. However, to ensure that the potential is there and can be met, we recommend the following:

1. That a survey be conducted of a) data bases available, b) data processes available, and c) user interest and mode of operation. In Tables 20 and 21 we have mocked up sample survey sheets on the data bases and processes. The third survey would be of a question-and-answer type designed to solicit both attitudes and information about current facilities that a given user might have for utilizing such a center. The surveyed users would represent the academic, government, and research communities.

2. That a pilot operation of an earth-physics information-management service be developed at a location where a diversity of earth-physics research talent and data-processing capabilities already exist. We recommend that the scope of this operation be no greater than the range anticipated in Table 18. The initial data bases and processes provided would be those indicated in this study.

3. That both the above projects be implemented in parallel, since the results of the survey could considerably augment the EPIMS data base but could still be inconclusive in terms of user attitude, which will develop with experience. Hence, we believe that actual operation of such a center will be the ultimate test of this concept.

Table 20.

EARTH-PHYSICS DATA-MANAGEMENT SURVEY

DATA BASE SHEET

NAME OF DATA BASE	*SAO Star Catalog*
PURPOSE OF DATA BASE	*Provide a full sky catalog of 258,997 stars with positions and proper motions for the epoch and equinox of 1950.0. All available catalogs were used provided they could be reduced to the FK4 frame of reference.*
STATUS	PROPOSED ☐ IN PROCESS ☐ OPERATIONAL ☒
DATES	ORIGINAL *1966* CURRENT VER. *1966* UPDATE FREQUENCY *None*

ORIGINATING ORGANIZATION		DISTRIBUTING ORGANIZATION	
NAME	*Smithsonian Astrophysical Observatory*	NAME	*Same*
ADDRESS	*60 Garden Street Cambridge, Mass. 02138*	ADDRESS	
CONTACT	*Mrs. K. Haramundanis*	CONTACT	

PRINCIPAL FIELD (please check one)

		GEOMETRICAL GEODESY ☐
GEOPOTENTIAL ☐	POSITIONAL ASTRONOMY ☒	EARTH KINETICS ☐
GEOLOGY ☐	OCEANOGRAPHY ☐	ATMOSPHERIC STUDIES ☐
IONOSPHERIC STUDIES ☐	MAGNETIC FIELD STUDIES ☐	INSTRUMENTS ☐
TIME & FREQUENCY ☐	OBSERVING STATIONS ☐	OTHER ______

RECORD CHARACTERISTICS — RECORD NAME: *Star*

FIELD	LENGTH	UNITS	FIELD	LENGTH	UNITS
A. *Magnitude, photo and visual*			F. *Mean epochs t_2 and t_2' of original observations*		
B. *Right ascension (α_{1950}) and declination (δ_{1950})*			G. *Annual proper motion for α and δ (μ and μ')*		
C. *Standard deviation (σ) at epoch 1950*			H. *Standard deviation of μ and μ'*		
D. *Right ascension (α_2) and declination (δ_2) from equator and equinox 1950.0*			I. *Spectral type*		
E. *Standard deviations (σ,σ') of α_2 and δ_2*			J. *Source catalog data*		

TOTAL NUMBER OF RECORDS *258,997* TOTAL CHARACTERS/RECORD *80*

STORAGE MEDIUM: PUNCH CARDS ☒ MAGNETIC TAPE 7 TRK ☒ 9 TRK ☐ OTHER ____

REMARKS AND REFERENCES	*Four versions of the Catalog exist on magnetic tape as well as the four-volume printed version and a corresponding set of 152 star charts at a scale of 120"/mm.*
FILLED OUT BY / DATE:	
NAME ORGANIZATION ADDRESS	

Table 21.

EARTH-PHYSICS DATA-MANAGEMENT SURVEY

DATA PROCESS SHEET

NAME OF DATA PROCESS	*Model Atmosphere*
PURPOSE OF DATA PROCESS	*To provide atmospheric pressure, temperature, and density as a function of altitude. Linear interpolation scheme is used to find the requested parameters between two stored values.*
STATUS	PROPOSED ☐ IN PROCESS ☐ OPERATIONAL ☒
DATES	ORIGINAL *Unknown* CURRENT VER. ____ UPDATE FREQUENCY *None*

ORIGINATING ORGANIZATION		DISTRIBUTING ORGANIZATION	
NAME	*North American Rockwell Corp.*	NAME	*Computer Software Management and Information Center (COSMIC)*
ADDRESS	*Canoga Park, California*	ADDRESS	*Barrow Hall, University of Georgia, Athens, Georgia*
CONTACT		CONTACT	

PRINCIPAL FIELD (please check one)

				GEOMETRICAL GEODESY	☐
GEOPOTENTIAL	☐	POSITIONAL ASTRONOMY	☐	EARTH KINETICS	☐
GEOLOGY	☐	OCEANOGRAPHY	☐	ATMOSPHERIC STUDIES	☒
IONOSPHERIC STUDIES	☐	MAGNETIC FIELD STUDIES	☐	INSTRUMENTS	☐
TIME & FREQUENCY	☐	OBSERVING STATIONS	☐	OTHER ____	

PROGRAM CHARACTERISTICS	PROGRAM NAME *Atmospheric Table Look-Up*
A. INPUT PARAMETERS	*Altitude*
B. OUTPUT PARAMETERS	*Pressure, temperature, density*
C. COMPUTER IMPLEMENTED ON	*IBM 7094*
D. CORE STORAGE	*14^K loc.*
E. TIME ESTIMATE	A. ELAPSED *20 sec/100 altitudes* B. CENTRAL PROCESSOR
F. LANGUAGE	*FORTRAN IV 97% map 3%*
G. NUMBER OF STATEMENTS	*2175*

STORAGE MEDIUM: PUNCH CARDS ☒ MAGNETIC TAPE 7 TRK ☒ 9 TRK ☐ OTHER ____

REMARKS AND REFERENCES		*Tables from Patrick A.F. Base Reference Atmosphere (0 to 155,000 ft.) and ARDC Model Atmosphere, 1959 (155,000 to 628,000 ft.). 100 altitudes maximum input. Option exists to print tables. Printout in metric or British system of units.*
FILLED OUT BY	DATE:	
NAME ORGANIZATION ADDRESS		

5. REFERENCES AND BIBLIOGRAPHY

ALLEN, T. J., GERSTENFELD, A., and GERSTBERGER, P. G.

1968. The problem of internal consulting in an R and D laboratory. Alfred P. Sloan School of Management Working Paper No. 319-68, Massachusetts Institute of Technology, Cambridge, Mass.

BERNSTEIN, W. A., and OWENS, J. T.

1968. Debugging in a time-sharing environment. Conf. Proc., Fall Joint Computer Conf., Thompson Book Co., Washington, D.C., vol. 33, pp. 7-14.

BULLARD, E. C., EVERETT, J. E., and SMITH, A. G.

1965. The fit of the continents around the Atlantic. Phil. Trans. Roy. Soc. London, vol. 258A, pp. 41-51.

CHAYES, F.

1963. Relative abundance of intermediate members of the oceanic basalt-trachyte association. Journ. Geophys. Res., vol. 68, pp. 1519-1534.

1964. Variance-covariance relations in some published Harker-diagrams of volcanic suites. Journ. Petrol., vol. 5, pp. 219-237.

1969. Statistical petrography. Carnegie Inst. Wash. Year Book, vol. 67, pp. 233-236.

CIRA

1965. <u>COSPAR International Reference Atmosphere 1965</u>. Compiled by COSPAR Working Group IV, North-Holland Publ. Co., Amsterdam, 313 pp.

COMPUTER INDUSTRY ANNUAL

1969. <u>Computer Industry Annual 1969-70</u>. Computerfiles, Inc., Concord, Mass.

CORBATO, F. J.

1968. Sensitive issues in the design of multi-use systems. MIT Project MAC Memo No. MAC-M-383, December.

CORBATO, F. J., and VYSSOTSKY, V. A.

1965. Introduction and overview of the MULTICS system. Conf. Proc., Fall Joint Computer Conf., Spartan Books, Washington, D.C., vol. 27, pp. 185-196.

DOLKOS, P. J.
1970. Multiplexing: Communications cost cutter. Data Processing Mag., vol. 12, pp. 34-40.

FRICKE, W., and KOPFF, A.
1963. Fourth Fundamental Catalogue. Veröff. Astron. Rechen-Inst., Heidelberg, No. 10, pp. 131-133.

FROHMAN, A.
1968. Determinants of library use in an industrial firm. Term paper, Alfred P. Sloan School of Management, Massachusetts Institute of Technology, Cambridge, Mass.

GAPOSCHKIN, E. M., and LAMBECK, K.
1970. 1969 Smithsonian Standard Earth (II). Smithsonian Astrophys. Obs. Spec. Rep. No. 315, 93 pp., May.

GRAHAM, R. M.
1968. Protection in an information processing utility. Comm. Assoc. Comp. Mach., vol. 11, pp. 365-369.

GURK, H. M., and MINKER, J.
1970. Storage requirements for information handling center. Journ. Assoc. Comp. Mach., vol. 17, pp. 65-77.

HARGRAVES, R. F., Jr., and STEPHENSON, A. G.
1969. Design considerations for an educational time-sharing system. Conf. Proc., Spring Joint Computer Conf., AFIPS Press, Montvale, New Jersey, vol. 34, pp. 657-664.

HARRISON, S., BROCKMAN, W. E., and SOMMERS, W. P.
1970. Analysis of large-scale aerospace data acquisition, processing, handling and dissemination systems. Private distribution.

HELA, I., and LISITZIN, E.
1967. A world mean sea level and marine geodesy. Proc. First Marine Geodesy Symp., U.S. Govt. Printing Office, Washington, D.C., pp. 71-73.

HOFMANN, R. B.
1968. Geodimeter fault movement investigations in California. Calif. Dept. Water Resources, Bull. 116-6, 183 pp.

HYMAN, H.
1967. The time-sharing business. Datamation, vol. 13, pp. 49-57.

JACCHIA, L. G.

1965. Static diffusion models of the upper atmosphere with empirical temperature profiles. Smithsonian Contr. Astrophys., vol. 8, pp. 215-257.

KAULA, W. M., Chairman

1970. The terrestrial environment: Solid earth and ocean physics. Report of a study at Williamstown, Mass., NASA Contractor Report NASA CR-1579, April.

KOZAI, Y.

1970. Seasonal variations of the geopotential. Smithsonian Astrophys. Obs. Spec. Rep. No. 312, 6 pp., April.

LAMPSON, B. W.

1969. Dynamic protection structures. Conf. Proc., Fall Joint Computer Conf., AFIPS Press, Montvale, New Jersey, vol. 35, pp. 27-38.

LeMAITRE, R. W.

1968. Chemical variation within and between volcanic rock series — a statistical approach. Journ. Petrol., vol. 9, pp. 220-252.

LETT, A. S., and KONIGSTORD, W. L.

1968. TSS/360: A time-shared operating system. Conf. Proc., Fall Joint Computer Conf., Thompson Book Co., Washington, D.C., vol. 33, pp. 15-28.

LUNDQUIST, C. A.

1971. Application of space techniques to solid-earth and ocean physics. In Space Research XI, Akademie-Verlag, Berlin, pp. 439-456, in press.

MARTIN, J.

1969. Telecommunications and the Computer. Prentice-Hall, Inc., New York, 470 pp.

MOHR, P. A.

1969. Relationships between recent Middle East and African earthquakes. Nature, vol. 223, pp. 516-518.

1970. Ethiopian rift and plateaus: some volcanic petrochemical distinctions. Submitted to Journ. Geophys. Res.

O'GARA, P. W.

1968. Physical location as a determinant of communication possibility among R and D engineers. S. M. Thesis, Alfred P. Sloan School of Management, Massachusetts Institute of Technology, Cambridge, Mass.

SACKMAN, H.

1968. Time-sharing versus batch processing: The experimental evidence. Conf. Proc., Spring Joint Computer Conf., Thompson Book Co., Washington, D.C., vol. 32, pp. 1-10.

SCHWARTZ, J. I.

1968. Interactive systems — promises, present and future. Conf. Proc., Fall Joint Computer Conf., Thompson Book Co., Washington, D.C., vol. 33, pp. 89-98.

SMITH, A. G., and HALLAM, A.

1970. The fit of the southern continents. Nature, vol. 225, pp. 139-144.

STAFF, SMITHSONIAN ASTROPHYSICAL OBSERVATORY

1966. Smithsonian Astrophysical Observatory Star Catalog. Smithsonian Institution, Washington, D.C.

STEEL, T. B., Jr.

1968. Multiprogramming — promise, performance and prospect. Conf. Proc., Fall Joint Computer Conf., Thompson Book Co., Washington, D.C., vol. 33, pp. 99-103.

TEPPER, M., Chairman

1969. Earth Surveys Program Documentation. Prepared by Earth Surveys Planning Panel, April 22.

TURCZYN, A., and JENG, T.

1970. Data communications: An overview of line and modern capabilities. Data Processing Mag., vol. 12, pp. 16-20.

U.S. STANDARD ATMOSPHERE

1962. U.S. Standard Atmosphere 1962. Prepared under the National Aeronautics and Space Administration, the United States Air Force, and the United States Weather Bureau, Washington, D.C., 278 pp.

U.S. STANDARD ATMOSPHERE SUPPLEMENTS

1967. U.S. Standard Atmosphere Supplements, 1966. U.S. Government Printing Office, Washington, D.C., 289 pp.

UTTERBACK, J. M.

1969. The process of technical innovation in industrial firms. Doctoral Thesis, Alfred P. Sloan School of Management, Massachusetts Institute of Technology, Cambridge, Mass.

VON ARX, W. S.

1967. Relationship between marine geodesy and oceanographic measurements. Proc. First Marine Geodesy Symp., U.S. Govt. Printing Office, Washington, D.C., pp. 37-42.

WEGNER, P.

1967. Machine organization for multiprogramming. Proc. 22nd National Conf., Assoc. Comp. Mach., Thompson Book Co., Washington, D.C., A.C.M. Publ. P-67, pp. 135-150.

WORLEY, A. R.

1969. Practical aspects of data communications. Datamation, vol. 15, pp. 60-66.